Chapter I
A Conceptual Framework for Data Mining and Knowledge Management

Shamsul I. Chowdhury
Roosevelt University, USA

ABSTRACT

Over the last decade data warehousing and data mining tools have evolved from research into a unique and popular applications, ranging from data warehousing and data mining for decision support to business intelligence and other kind of applications. The chapter presents and discusses data warehousing methodologies along with the main components of data mining tools and technologies and how they all could be integrated together for knowledge management in a broader sense. Knowledge management refers to the set of processes developed in an organization to create, extract, transfer, store and apply knowledge. The chapter also focuses on how data mining tools and technologies could be used in extracting knowledge from large databases or data warehouses. Knowledge management increases the ability of an organization to learn from its environment and to incorporate knowledge into the business processes by adapting to new tools and technologies. Knowledge management is also about the reusability of the knowledge that is being extracted and stored in the knowledge base. One way to improve the reusability is to use this knowledge base as front-ends to case-based reasoning (CBR) applications. The chapter further focuses on the reusability issues of knowledge management and presents an integrated framework for knowledge management by combining data mining (DM) tools and technologies with CBR methodologies. The purpose of the integrated framework is to discover, validate, retain, reuse and share knowledge in an organization with its internal users as well as its external users. The framework is independent of application domain and would be suitable for uses in areas, such as data mining and knowledge management in e-government.

INTRODUCTION

People have been collecting and organizing data from stone ages. In the earlier days data were collected and recorded in one way or the other mainly for record keeping purposes. With the advancement in computational technology in general and storage technology in particular data collection and their storage in large data warehouses have become an integral part of the data processing and decision-making environment of today's organizations. Over time people have learned to value data as an important asset.

Reliable data in a database or a data warehouse could be used for decision-making purposes by appropriately analyzing the data and making them more meaningful and useful. In other words data could be analyzed to find hidden patterns and foresee trends. The process is broadly being called data mining.

Data mining usually starts with a hypothesis or an assumption and ultimately creates new information or knowledge. In order to survive and succeed in the tough business world of today it is also very important to store and manage the evolving knowledge within an organization. Knowledge management is the overall activities of creating, storing, re-using and sharing the new knowledge. Knowledge management increases the ability of an organization to learn from its environment and to incorporate knowledge into the business processes by adapting to new tools and technologies, for example; data warehousing, data mining and case-based reasoning.

The chapter presents the necessary fundamentals of data warehousing (DW), data mining (DM) (methodology, tools, techniques, systems and terminology) and related technologies. One of the purposes of the chapter is to develop and gain an understanding of the principles, concepts, functions and uses of data warehousing and data mining for knowledge management in a broader sense. Data modeling in data warehousing plays a vital role in successful utilization of the data

resources as an organizational asset. The data quality is also an important aspect in the process. The following topics have been addressed:

*   DW methodology
*   DW data modeling
*   DW data quality issues
*   DM – techniques and uses
*   DM and Business intelligence – From findings to application
*   DM and CBR (Case-based Reasoning) for knowledge retention and reuse.

The work falls into the category of content management of data from database/data warehouse using data mining and other intelligent techniques like expert system, CBR, etc. The purpose is to make the extracted knowledge available to the users (both internal and external) in organizations of different structure and forms, including e-government. In e-government one of the most important benefits would be citizen empowerment through access to information/knowledge (The World Bank Group 2008).

In the work the term *business* and *customer* has been used in a broader sense. By *business* and *customer* is meant the core activities that take place and the consumer of information respectively in an organization.

BACKGROUND

Over the years data warehousing and data mining tools have evolved into a unique and popular business solutions to attain business understanding and decision-making. Decision makers already consider these systems to be the corner stone in their IT system portfolio. Data and the knowledge derived from the data and sharing them within an enterprise and its business partners and collaborators is a key success factor in today's complex business world, and hence flexible tools are needed to deal with the evolving complexities.

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/conceptual-framework-data-mining-knowledge/29061

Related Content

Mining Allocating Patterns in Investment Portfolios
Yanbo J. Wang, Xinwei Zhengand Frans Coenen (2009). *Data Mining Applications for Empowering Knowledge Societies (pp. 110-135).*
www.irma-international.org/chapter/mining-allocating-patterns-investment-portfolios/7549

Hybrid Recommender System Using Emotional Fingerprints Model
Anthony Nosshi, Aziza Saad Asemand Mohammed Badr Senousy (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines (pp. 1076-1100).*
www.irma-international.org/chapter/hybrid-recommender-system-using-emotional-fingerprints-model/308534

A Framework on Data Mining on Uncertain Data with Related Research Issues in Service Industry
Edward Hung (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications (pp. 515-529).*
www.irma-international.org/chapter/framework-data-mining-uncertain-data/73455

Communicating Natural Calamity: The Sentiment Analysis of Post Rigopiano's Accident
Nicola Capolupoand Gabriella Piscopo (2022). *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines (pp. 1528-1539).*
www.irma-international.org/chapter/communicating-natural-calamity/308559

Coupling Materialized View Selection to Multi Query Optimization: Hyper Graph Approach
Ahcene Boukorca, Ladjel Bellatreche, Sid-Ahmed Benali Senouciand Zoé Faget (2015). *International Journal of Data Warehousing and Mining (pp. 62-84).*
www.irma-international.org/article/coupling-materialized-view-selection-to-multi-query-optimization/125651